

CORRECCIÓN POR SESGOS EN MODELOS LOG-NORMALES ALOMÉTRICOS LINEALIZADOS UTILIZADOS PARA LA ESTIMACIÓN DE LA BIOMASA AÉREA

Mauricio Zapata¹, Jorge Ignacio del Valle² y Sergio Alonso Orrego³

¹Estudiante Ingeniería Forestal, Universidad Nacional de Colombia, Sede Medellín.
e-mail: mauricio_zapata@colforest.com.co

²Profesor titular, Departamento de Ciencias Forestales, U. Nacional de Colombia, Sede Medellín.
e-mail: jidvalle@perseus.unalmed.edu.co

³Profesor titular, Departamento de Ciencias Forestales, U. Nacional de Colombia, Sede Medellín.
e-mail: saorrego@perseus.unalmed.edu.co

RESUMEN

El compartimiento más importante en proyectos de captura de carbono, tanto en peso como en su tasa de cambio es la biomasa aérea. Para los bosques del área de Porce, Colombia, los árboles representan 89 % de la biomasa aérea total. En coberturas boscosas se ha generalizado la estimación de la biomasa aérea del árbol por medio de una relación estadística de tipo alométrica, cuya variable independiente es el diámetro normal. Siempre se ha encontrado que esta relación presenta una varianza condicional heterogénea, por lo cual se hacen las transformaciones logarítmicas de las variables para realizar el análisis de regresión lineal con varianza homogénea.

El empleo de estas transformaciones introduce un sesgo sistemático en las predicciones cuando se revierten a la forma original no lineal y, por tanto, estimaciones sesgadas. En la literatura no se corrigen estos sesgos, o no se tienen en cuenta.

En este trabajo se utilizó la información de 32 parcelas permanentes para evaluar el impacto en la estimación del carbono almacenado en 609 ha de bosques primarios, al usar un modelo de biomasa aérea alométrico con y sin corrección por sesgos. Se evaluaron cuatro propuestas de corrección de las estimaciones obtenidas con el modelo: la primera consiste en no efectuar ninguna corrección; la segunda utiliza el Cuadrado Medio del Error del modelo lineal log-normal; la tercera es un factor derivado de funciones de Bessel modificadas que se aproximan a términos de orden $1/n^2$ para obtener las estimaciones insesgadas del modelo de regresión; la cuarta es un factor que trabaja con la solución de Finney para estimaciones insesgadas y de mínima varianza de una variable log-normal adaptada a modelos de regresión.

La omisión del factor de corrección al modelo alométrico genera subestimaciones en la cantidad de biomasa y, por tanto, de carbono almacenado (t/ha) del orden de 24%. Esto equivale, en términos de créditos de carbono, a una subestimación de 15.374 t.

Palabras claves: Colombia, biomasa aérea, corrección por sesgos, ecuaciones de biomasa, secuestro de carbono, modelos alométricos.

Introducción

La estimación de la biomasa de los bosques tropicales es un tema relevante en relación con el problema del calentamiento global. Por esto, recientemente se ha incrementado el interés por estudiar el papel de los bosques en los ciclos de elementos biogeoquímicos, especialmente del carbono (Delaney *et al.* 1997), y su relación con los gases con efecto de invernadero (Brown 1997). Se parte del principio de que aproximadamente 50% de la biomasa estimada es carbono (Brown & Lugo 1992, Brown 1997) y, por tanto, la biomasa del bosque representa la cantidad de carbono que puede ser adicionada a la atmósfera como dióxido de carbono (CO₂), cuando éste se corta y se quema. El exceso de CO₂ en la atmósfera acentúa el efecto invernadero. Ello trae como consecuencia cambios en variables climáticas globales (Morrissey & Justus 1998).

En el ámbito internacional se han iniciado estrategias para el control del calentamiento global como la negociación de créditos de carbono. Para hacer realidad estas estrategias en proyectos forestales es necesario medir y monitorear el carbono almacenado, especialmente en el compartimiento aéreo que es el más grande y fácil de manejar (MacDicken 1997). Las ecuaciones de biomasa son la mejor aproximación para lograr estas estimaciones y, como consecuencia, se adopta en casi todas las investigaciones relacionadas con la cuantificación de la biomasa de los bosques tropicales (Saldarriaga *et al.* 1988, Higuchi *et al.* 1998).

Al graficar una cantidad de datos representativos de diámetro normal de los árboles contra su masa seca se observa una tendencia no lineal, indicando una relación alométrica de la forma (Overman *et al.* 1994, Enquist *et al.* 1998)

$$W = aD^bV, \quad (1)$$

donde W es la masa seca del árbol en kg, D es el diámetro del árbol a 1,3 m sobre el suelo en cm, a y b son parámetros del modelo y V es el término multiplicativo del error.

Sin embargo, la nube de puntos muestra un problema de heterocedasticidad que corresponde a un incremento de la varianza con el aumento del diámetro (Overman *et al.* 1990,1994, Álvarez 1993). La mayoría de los trabajos en biomasa utilizan el método inicialmente propuesto por Kittredge (1944). Éste consiste en que la masa seca de algún componente, o del árbol, se relaciona con variables de tamaño del árbol en una regresión lineal bajo escalas logarítmicas. Esta transformación simplifica los cálculos, pues permite usar análisis de regresión lineal y, además, incrementa la validación estadística del análisis al homogeneizar la varianza sobre el rango de los datos. Este procedimiento es el más recomendado para estimar la biomasa en los bosques tropicales (Brunig 1983, Scatena *et al.* 1993, Chave *et al.* 2001).

Algunos autores (Baskerville 1965, Jordan & Uhl 1978, Crow 1980, Sprugel 1983, Saldarriaga *et al.* 1988, Overman *et al.* 1990 Scatena *et al.* 1993, Regina 2000, Chave *et al.* 2001) han utilizado la relación alométrica linealizada con transformaciones logarítmicas para tal fin. Cuando se elige hacer transformaciones de los datos se está asumiendo que los términos del error tienen una distribución log-normal; es decir, que los logaritmos de los errores se distribuyen normalmente (Heien 1968, Bradu & Mundlak 1970, Carroll & Ruppert 1988), de tal modo que los v ($\ln(V)$) satisfacen los supuestos estadísticos.

Sin embargo, cuando se recupera el modelo original tomando antilogaritmos, se introduce un sesgo sistemático en las predicciones que da lugar a estimaciones sesgadas (Heien 1968, Thöni 1969, Bradu & Mundlak 1970, Beauchamp & Olson 1973, Mountford & Bunce 1973, Sprugel 1983).

En la literatura se encuentran varias propuestas para realizar las correcciones por sesgos a los parámetros estimados de modelos con transformaciones logarítmicas. Por ejemplo, Finney en 1941 inicialmente encontró estimadores insesgados y de mínima varianza de una variable aleatoria que se distribuye como log-normal, incorporando correcciones con series infinitas (Bradu & Mundlak 1970). Laurent (1963) trabajó con funciones de Bessel modificadas para encontrar estimadores de mínima varianza para la media y la varianza. Beauchamp & Olson (1973) trabajaron con aproximaciones a términos de orden $1/n^2$ para las estimaciones de la varianza en los modelos de regresión.

En este trabajo se analizan las propuestas de corrección por CME (Cuadrado Medio del Error), corrección de Finney, adaptadas a modelos de regresión por Bradu & Mundlak (1970) y las de Beauchamp & Olson (1973), con la idea de evaluar el efecto que tienen sobre las estimaciones de la biomasa aérea de los modelos alométricos ajustados linealmente, y analizar el impacto en proyectos de captura de carbono en bosques tropicales.

Métodos

ÁREA DE ESTUDIO.— Los bosques primarios tropicales donde se realizó el estudio de biomasa hacen parte del área de influencia de la Central Hidroeléctrica Porce II, localizada en la cuenca media del río Porce, cordillera Central de los Andes Colombianos (Figura 1).

Las parcelas permanentes se encuentran en un rango de elevación de 925 m a 1500 m. La precipitación media anual es 1977 mm. La temperatura promedio es 22,8° C, con valores promedios mínimos de 17,9° C y máximos de 29,3° C (EPM 1992). En las 609 ha de bosques primarios se han establecido 32 parcelas permanentes de 0,1 ha para un proyecto de monitoreo de captura de carbono por las coberturas vegetales, y donde se monitorea el incremento radial de los árboles con diámetros mayores de 10 cm y el incremento de diámetro para los árboles entre 1 cm y 10 cm en subparcelas de 10 m x 10 m (Orrego & del Valle 2001).

Se utilizará la información del inventario del año 2001 para evaluar el efecto de las correcciones por sesgos en las estimaciones de las existencias de biomasa y, por tanto, de carbono almacenado en 609 ha de bosques primarios.

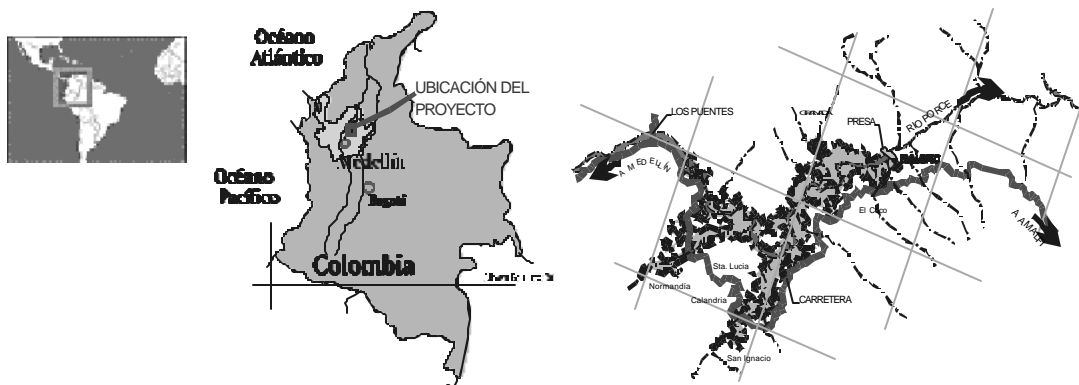


FIGURA 1. Ubicación de la Central Hidroeléctrica Porce II.

MÉTODOS DE CAMPO Y LABORATORIO.— Los árboles utilizados para las ecuaciones se seleccionaron con el criterio de su diámetro normal de modo aleatorio en todos los tamaños. El diámetro mínimo fue de 0,5 cm y el máximo determinado por el árbol de mayor diámetro hallado, 198 cm. A cada árbol seleccionado se le midió con cinta métrica el perímetro del fuste a 1,3 m sobre el suelo. En los árboles con bambas su perímetro se tomó en un sitio más arriba de donde éstas se desprenden y donde el fuste era más cilíndrico.

Los árboles se apearon haciendo un corte a ras del suelo. Una vez apeado se midió la longitud del fuste con cinta métrica. Para encontrar la masa verde del árbol se trabajó por componentes (hojas, ramitas con diámetros menores de 2,5 cm, ramas y fuste). Los árboles con diámetros de fuste superiores a 30 cm de diámetro normal se cubicaron. El peso seco de los elementos cubicados se encontró posteriormente con base en muestras tomadas para la determinación de la densidad básica.

Para determinar el peso seco de los componentes en cada árbol, y hallar su contenido de humedad, se tomaron muestras representativas de 500 g para cada componente. A cada árbol se le recolectaron, además, dos muestras botánicas para su posterior identificación. Todo el material de las muestras se secó en horno a $103^{\circ} \pm 2^{\circ}$ C hasta alcanzar peso constante. La masa seca de cada componente, por árbol, se determinó con la relación masa seca a masa verde de las muestras en cada componente. La masa seca de cada árbol se obtuvo al sumar todas las masas secas de los componentes de cada árbol.

MÉTODOS DE ESTIMACIÓN.— Para la ecuación de biomasa se estimaron los parámetros de un modelo alométrico linealizado con transformación logarítmica de las variables, usando estimadores de máxima verosimilitud. El modelo es de la siguiente forma (Zapata 2001)

$$y = \ln a + b \ln D + \ln V, \quad (2)$$

donde $y = \ln W$, es el logaritmo de la masa seca en kg de un árbol de diámetro D . La razón de esta propuesta subyace en las evidencias de que en el tipo de relaciones de tamaños, por ejemplo el de la ecuación (1), funciona los supuestos hechos con respecto al tipo de función de distribución que siguen las variables. Esta característica, considerada antes de las estimaciones de los parámetros, hace que estos sean eficientes. Esto no sucede con los estimadores mínimo cuadráticos y más aún, cuando la $\text{var}(V_i) = D_i^\alpha \sigma^2$; es decir, la varianza es proporcional a una potencia de la variable independiente D . En este caso los estimadores mínimo cuadráticos son ineficientes (Judge *et al.* 1987, Carroll & Ruppert 1988, Green & Valentine 1998).

CORRECCIÓN POR SESGOS.— Al modelo ajustado para la biomasa aérea se le aplicaron cuatro correcciones por sesgos, debidas a la transformación logarítmica, y se compararon los resultados de sus predicciones. En todos los trabajos de biomasa reportados se ha aplicado alguna de las cuatro correcciones propuestas. Estas son: (1) no hacer corrección por sesgos; (2) usar el factor $(\text{CME}/2)$ para el parámetro constante; (3) utilizar las aproximaciones de orden $1/n^2$ propuestas por Beauchamp & Olson (1973) para obtener estimaciones insesgadas de las predicciones, y (4) utilizar la solución de Finney para modelos de regresión propuesta por Bradu & Mundlak (1970).

CORRECCIÓN POR **CME**.— El problema radica en estimar parámetros insesgados de una función que con las transformaciones logarítmicas resulta en una expresión lineal.

Para un conjunto de n observaciones de W y D se puede escribir la relación estadística como una modificación de la conocida función Cobb-Douglas

$$W = a_0 D_i^{b_1} V, \quad (i = 1, \dots, n) \quad (3)$$

donde W representa la biomasa aérea de los árboles en kg, D es su diámetro normal en cm, V es el término multiplicativo del error que tiene una distribución lognormal (Heien 1968).

DERIVACIÓN DEL TÉRMINO DE CORRECCIÓN.— Si V_1, V_2, \dots, V_n son los parámetros aleatorios independientes del error y

$$\mathbf{v} = \prod_{i=1}^n V_i, \quad (4)$$

entonces

$$\ln \mathbf{V} = \sum_{i=1}^n \ln V_i. \quad (5)$$

Si a los parámetros aleatorios independientes $\ln V_i$ se les aplicara el resultado del teorema central del límite, entonces la distribución estandarizada de los $\ln \mathbf{V}$ tiende a una normal cuando n tiende a infinito, y la distribución límite de los \mathbf{V} tiende a una lognormal biparamétrica.

Transformando la ecuación (3) al tomar logaritmos naturales en ambos lados de la función se obtiene

$$y_i = \beta_0 + \beta_1 x_i + v_i, \quad (6)$$

donde $y = \ln W_i$, $x = \ln D_i$, $\beta_0 = \ln a_0$, y $v = \ln \mathbf{V}_i$. El procedimiento acostumbrado es ajustar la ecuación (6) para obtener el vector de estimadores β (β_0 y β_1) por métodos de regresión lineal. Con las propiedades de V se puede estimar el conjunto de parámetros β por métodos de máxima verosimilitud, partiendo de que cada v se distribuye $N(0, \sigma^2)$.

El primer momento de tendencia de la distribución lognormal es conocido y relaciona la media de una variable lognormal (ζ) con la media (m) y la varianza (σ^2) de una variable aparentemente normal. Esto es

$$\hat{\zeta}_r = E(V^r) = \text{EXP} \left(rm + \frac{1}{2} r^2 \sigma^2 \right). \quad (7)$$

Para el primer momento de tendencia central $r=1$,

$$\bar{x} = \hat{\zeta}_1 = \text{EXP} \left(m + \frac{1}{2} \sigma^2 \right), \quad (8)$$

dado que la media (m) de v es cero la media de \mathbf{V} será $\bar{x} = \text{EXP}\left(\frac{1}{2}\hat{\sigma}^2\right)$. Bajo estos supuestos es claro que (Heien 1968)

$$E(W_i) = a_0 D_i^{b_1} \text{EXP}\left(\frac{1}{2}\hat{\sigma}^2\right) \neq a_0 D_i^{b_1} \quad (i=1, \dots, n) \quad (9)$$

REPARAMETRIZACIÓN DEL MODELO.— Al describir el modelo (3) con un nuevo término multiplicativo U se tiene

$$W = a_0 D_i^{b_1} \cdot U \quad (i=1, \dots, n), \quad (10)$$

de modo que el modelo requerido es

$$E(W_i) = a_0 D_i^{b_1} \quad (i=1, \dots, n). \quad (11)$$

Lo anterior significa que $E(U_i) = 1$, ($i=1, \dots, n$), y para el modelo con transformaciones logarítmicas

$$y_i = \beta_0 + \beta_1 x_i + u_i, \quad (12)$$

en donde el vector del término del error es $N(-1/2\hat{\sigma}^2, \hat{\sigma}^2)$. El fundamento de su distribución está dado por (8), para que con las transformaciones logarítmicas de (10) u ($\ln U_i$) tenga un valor esperado de cero. Rescribiendo (12) se tiene

$$y_i = \hat{\beta}_0' + \beta_1 x_i + v_i \quad (i=1, \dots, n), \quad (13)$$

donde $\hat{\beta}_0' = \hat{\beta}_0 - (1/2)\hat{\sigma}^2$ y v es $N(0, \hat{\sigma}^2)$. Si a $\hat{\beta}_0$ se le resta el valor medio de u se puede garantizar que la variable que ocupa el lugar del término del error tiene una media cero.

En el modelo (13) se puede encontrar el conjunto de parámetros estimados β mediante análisis de regresión. Para el mejor conjunto de parámetros se cumple que $E\left(\hat{\beta}_0'\right) = \beta_0'$, $E\left(\hat{\beta}_1\right) = \beta_1$ y

$E(S^2) = \sigma^2$; de lo anterior resulta evidente que $E\left[\hat{\alpha}_0' + (S^2/2)\right] = \hat{\alpha}_0' + (1/2)\sigma^2 = \hat{\alpha}_0$, y $\beta_0 = \hat{\alpha}_0' + (1/2)\sigma^2$ es un estimador insesgado de $\ln(a)$

SOLUCIÓN DE FINNEY.— Ahora se considera la solución de Finney (Bradú & Mundlak 1970) para encontrar los estimadores insesgados y de mínima varianza para una variable W (variable log-normal).

Sea $Z_i = \ln(W)$, $S^2 = (n-1)^{-1} \sum_{i=1}^n (Z_i - \bar{Z})^2$ y $\bar{Z} = n^{-1} \sum_{i=1}^n Z_i$. Si una función $f(\bar{Z}, S^2)$ tiene un valor esperado $h(\zeta, \sigma)$ bajo la nueva distribución (después de la transformación), éstos estadísticos (\bar{Z}, S^2) son estimadores insesgados y de mínima varianza de $h(\zeta, \sigma)$. A su vez, el valor esperado de $\text{EXP}(\zeta + 1/2\sigma^2)$ es \bar{W} (de acuerdo con lo explicado arriba), y el valor esperado de $e^{\sigma^2} (e^{\sigma^2} - 1)e^{2\zeta}$ es la varianza de la variable W . Finney obtiene los estimadores del valor esperado y de la varianza (V) de infinitas series así

$$\bar{W} = \text{EXP}(\bar{Z}) \cdot g\left(\frac{1}{2}S^2\right), \quad (14)$$

$$V = \text{EXP}(2\bar{Z}) \{g(2S^2) - g((n-2)S^2/(n-1))\}, \quad (15)$$

donde

$$g(t) = 1 + \frac{n-1}{n}t + \sum_{j=2}^{\infty} \frac{(n-1)^{2j-1}}{n^j (n+1)(n+3)\cdots(n+2j-1)} \frac{t^j}{j!}. \quad (16)$$

Para el modelo de regresión se tiene (Bradú & Mundlak 1970)

$$\bar{W} = \text{EXP}(\hat{n}) \cdot g\left[\frac{m+1}{2m}(\sigma^2 - \sigma^2_{\hat{n}})\right] \quad (17)$$

donde \hat{n} es el modelo linealizado, $m = n-k$ (parámetros) y $\sigma^2_{\hat{n}}$ es la varianza de una nueva observación.

CORRECCIÓN POR TÉRMINOS DE ORDEN $1/n^2$.— Un estimador de mínima varianza insesgado de W es $a_1 = (\hat{w}, \hat{\sigma}^2) = \text{EXP}(\hat{a}_0 + \hat{a}_1 x) \theta(\hat{\sigma}^2/2)$, donde $\psi(\sigma^2/2)$ es una función de Bessel modificada que aproximada en términos de $1/n^2$ es (Beauchamp & Olson 1973)

$$\psi(\sigma^2/2) = \left[1 - \frac{\sigma^2(2\phi/n)}{4n} + \frac{(\sigma^2)^2 \left[(\sigma^2)^2 + 2(16/3 + 2\phi)\sigma^2 + 4\phi^2 + 16\phi \right]}{32n^2} \right], \quad (18)$$

donde $f = \sum_{i=1}^m (w_i - w)^2 / \sum_{i=1}^n (w_i - \bar{w})^2$.

PRECISIÓN DE LAS ESTIMACIONES.— La precisión de los modelos se evaluó con las propuestas de Overman *et al.* (1994), pero con una modificación como sigue

$$\delta B = \frac{\sum_{i=1}^n \frac{(\hat{D}W_i - DW_i) * 100}{DW_i}}{n}, \quad (19)$$

donde $\hat{D}W_i$ es la masa estimada del árbol i y DW_i es la masa observada del árbol i . Entre más pequeño es el factor, más preciso es el modelo. El porcentaje obtenido refleja el valor correcto de sobrestimación o subestimación del modelo

Resultados

La muestra de árboles medidos, cosechados y pesados en los bosques primarios de Porce, para la estimación de las ecuaciones de biomasa aérea, está constituida por 144 árboles distribuidos en un rango diamétrico de 0,28 cm a 198,8 cm. En la Figura 2 se presenta la distribución de los árboles por clases de tamaño de 5 cm. Se observa que la forma de la distribución de los árboles por clases de tamaño sigue una J invertida, efecto explicado por la facilidad de tomar información en las clases de tamaño inferiores.

En la Tabla 1 se presentan las principales características de las variables que se utilizaron en la estimación de las ecuaciones de biomasa. Los datos se encuentran en un rango de biomasa (Bt) entre 0,07 kg para un *Jacaranda copaia* y 27.504,8 kg para un caimo (*Manilkara* sp.). La biomasa del árbol promedio en la muestra es de 1572,69 kg. La biomasa total de la muestra de árboles pesada en campo es de 226.467 kg. Esta masa representa 72,23% de la biomasa promedio de los

árboles en una hectárea de bosques húmedos tropicales (313.518 kg/ha, Zapata 2001). Esto da una idea de que la muestra de árboles es representativa en masa con respecto al promedio mundial.

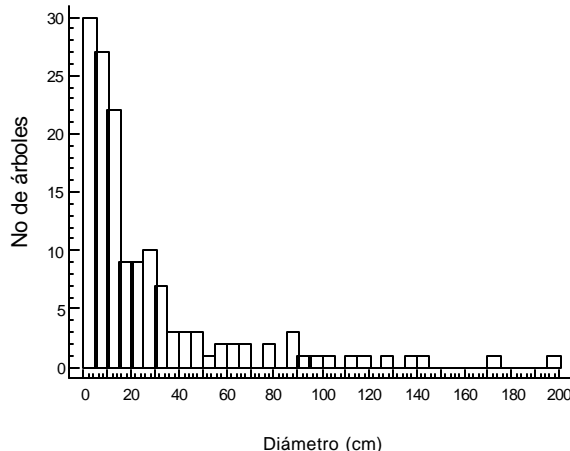


FIGURA 2. Distribución de la muestra de árboles por clases de tamaño del diámetro normal (D).

El menor valor de la variable diámetro (D) es de 0,28 cm, que corresponde a un individuo de la familia Lauraceae y el diámetro más alto es de 198,88 cm que corresponde a un caracolí (*Anacardium excelsum*). Este caracolí aunque posee el mayor tamaño en diámetro no es el árbol con mayor biomasa, debido a que la densidad de su madera ($0,441 \text{ g/cm}^3$) es menor que la de un caimo (*Manilkara* sp.) que tiene 174,12 cm de diámetro y es el más pesado ($0,744 \text{ g/cm}^3$). El valor del diámetro del árbol medio de la muestra es de 27,17 cm.

TABLA 1. Características de las variables tomadas en la muestra de árboles cosechados para estimar las ecuaciones de biomasa.

Variable	Rango		\bar{X}	S^a	Total	N
	Inferior	Superior				
Bt (kg)	0,07	27504,8	1572,69	4068,81	226467,0	144,0
D (cm)	0,28	198,88	27,1728	35,2116	--	144,0
H (m)	1,26	46,7	15,6394	11,1734	--	144,0

^a Desviación estándar

La Tabla 2 contiene los resultados de la estimación de los parámetros máximo verosímiles de un modelo alométrico linealizado con transformaciones logarítmicas, al igual que algunos estadísticos

de ajuste y precisión. En la Figura 3 se muestra la línea ajustada por el modelo alométrico tanto en escala logarítmica como en la escala original. La relación entre la biomasa total de los árboles y la variable diámetro (D) sigue el mismo patrón indicado por Bell *et al.* (1984), en forma de cola de cometa (Figura 3).

TABLA 2 *Parámetros estimados del modelo alométrico con sus estadísticos más importantes. E.E. error estándar de parámetros, N número de árboles, s desviación estándar del error, I razón de máxima verosimilitud, R² coeficiente de determinación, dB porcentaje medio de desviación absoluta modificada.*

	Modelo	E. E.		N	s	I ^a	R ²	dB ^b
		a	b					
1	$\ln(Bt) = -2,2862 + 2,4709 \ln(D)$	0,0890	0,0304	140	0,426	544,1 0	0,979	7,70

^a significativo para valores mayores de $\chi^2_{\alpha=0,05}=5,99$

^b Según la ecuación (19).

EL modelo linealizado se ajustó con 140 datos, debido a que se presentaron 4 observaciones con más de tres desviaciones estándar en sus residuales estandarizados que fueron excluidos.

El valor del parámetro dB indica que el modelo hallado para estimar la biomasa aérea total, está sobrestimando en promedio la biomasa de los árboles de la muestra en 7,7 %. En la Tabla 3 también se utiliza el porcentaje de desviación dB para evaluar los efectos de los cuatro factores de corrección por sesgos sobre las estimaciones.

La propuesta de corrección (2) es la única que permite escribir de nuevo un modelo insesgado, pues adopta un factor constante para la corrección ($CME/2$). Los factores de corrección (3) y (4) adoptan valores específicos de corrección dependiendo del nivel de la variable D ; por esta razón no se presenta la magnitud de corrección en la Tabla 3. Como era de esperarse, la aplicación de cualquiera de los factores, excepto el (1), desplaza las líneas de regresión por encima de sus valores sesgados (Figura 4).

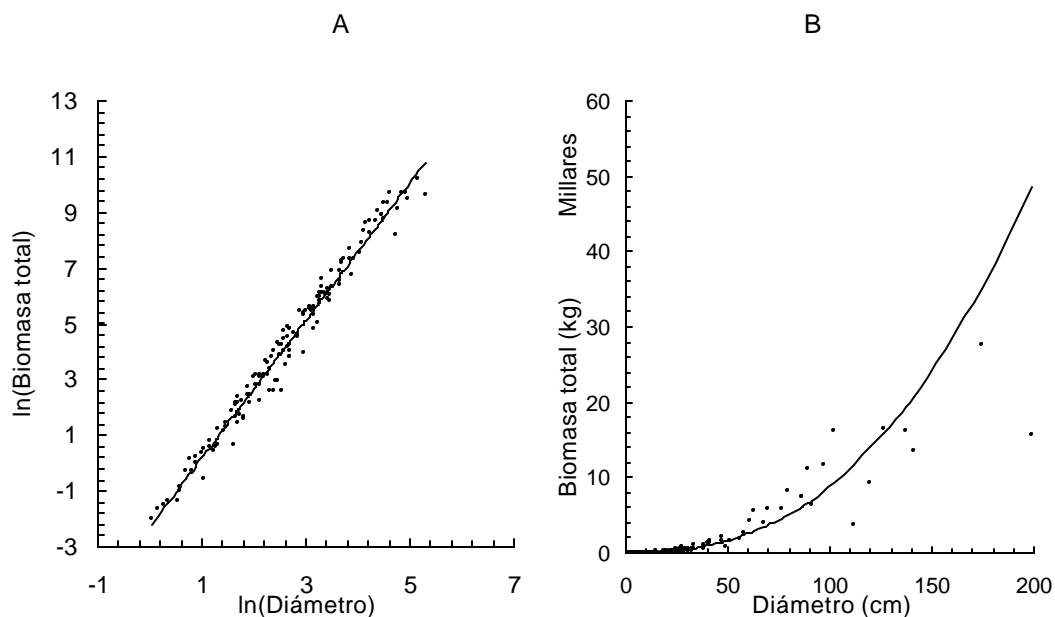


FIGURA 3. A: líneas ajustada de regresión del modelo alométrico para la biomasa total de los árboles con el diámetro normal. B: relación de la biomasa total de los árboles y su diámetro para el modelo alométrico en su escala original.

También se puede observar en la Figura 4 que las líneas de regresión corregidas por los diferentes factores se superponen indicando la similitud en sus predicciones.

TABLA 3. Evaluación del efecto de las correcciones por sesgos debidas a las transformaciones logarítmicas en el modelo de biomasa. El rango hace referencia a los valores de diámetro usados para la estimación del modelo. N es el número de árboles utilizados en el modelo de regresión. $d\mathcal{B}_i$ porcentaje de desviación obtenida con el factor de corrección i . Factor (i) es la magnitud de corrección del modelo.

Rango (cm)	Modelo	N	$d\mathcal{B}_1$	Factor (2)	$d\mathcal{B}_2$	Factor (3)	$d\mathcal{B}_3$	Factor (4)	$d\mathcal{B}_4$
0,28-200	$0,1017 D^{2,4709}$	140	7,70	0,09139	21,37		21,257		21,259

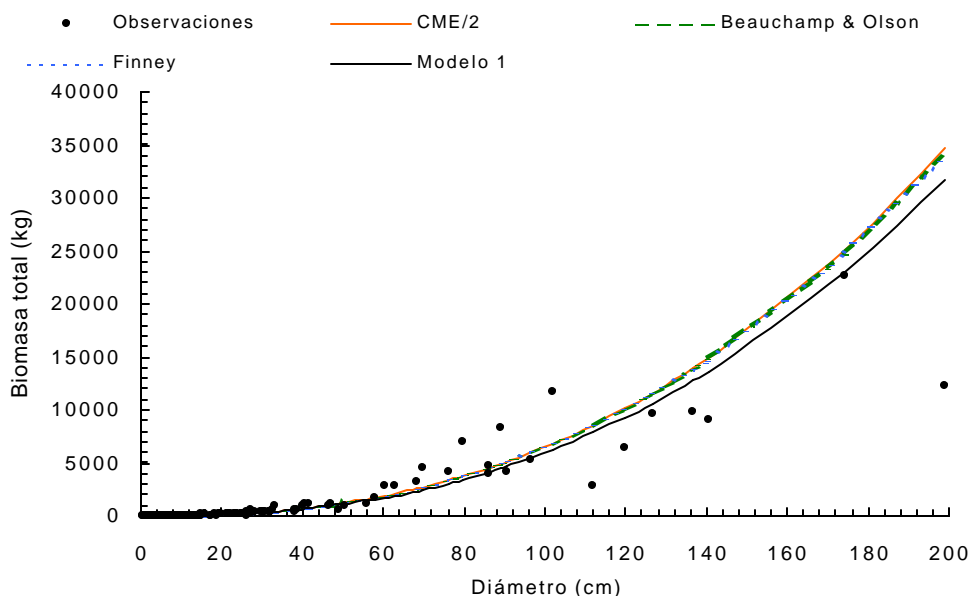


FIGURA 4. Representación de las líneas de modelos alométricos insesgados según las propuestas de corrección de la Tabla 3. Se comparan con el modelo para la biomasa total aérea sin corrección $Bt = 0,1017 D^{2,4709}$.

Utilizando el modelo alométrico de la Tabla 2 y aplicando los diferentes factores de corrección sobre las estimaciones de biomasa para la información de 32 parcelas permanentes establecidas en Porce, se obtiene que el impacto de la corrección promedio es de 24 % con respecto a la obtenida con el factor (1) (Tabla 4). Para el área en bosques primarios del proyecto (609 ha) la omisión de un factor como el (2), que es el más sencillo, produce una subestimación del carbono almacenado de 15.373,90 t.

TABLA 4 Evaluación del impacto de las correcciones por sesgos en la estimación de la biomasa aérea usando 32 parcelas permanentes establecidas en Porce (Colombia).

Factor de corrección	Biomasa estimada (t/ha)	Porcentaje de corrección	Impacto de la corrección en términos de carbono (t/ha)	Impacto sobre el área en bosques primarios del proyecto en términos de carbono (t)
Factor (1)	211,038	0	0	0
Factor (2)	261,527	23,92	25,24	15.373,90
Factor (3)	260,512	23,44	24,74	15.064,83
Factor (4)	260,519	23,45	24,74	15.066,96
Promedio (2 a 4)	260,853	23,60	24,91	15.168,566

Discusión

Muchos estudios ecológicos han usado modelos de regresión logarítmica por mucho tiempo para estimar varios componentes biológicos, pero ponen poca atención al problema de los sesgos al hacer las predicciones, y en los errores al usar estos modelos (Mountford & Bunce 1973). Prueba de esto lo constituye el reporte en la literatura de modelos para estimar la biomasa aérea total de los árboles de bosques tropicales del mundo ajustados con transformaciones logarítmicas, pero sin hacer consideraciones en relación con los sesgos (Tabla 5).

TABLA 5. Resumen de modelos alométricos reportados para estimar la biomasa aérea total de algunos bosques tropicales. La biomasa aérea del árbol se representa con y en kg, D es el diámetro normal en cm. MHT = Muy Húmedo Tropical, HT = Húmedo Tropical, LT = Lluvioso Tropical, HST = Húmedo Subtropical, MHT = Muy Húmedo Tropical.

Localidad	a	b	R^2	N	Rango	Tipo de bosque	Fuente
Modelo: $\ln y = \ln a + b \ln D$							
Brasil	-2,43	2,57	0,96	30	--	--	Brown (1997) ^a
Bisley, Luquillo, Puerto Rico	-2,399	2,475	0,926	63	2,5-57	HST	Scatena <i>et al.</i> (1993)
Bosques tropicales	-2,14	2,41	0,98	146	>10	≥3000mm	Chave <i>et al.</i> (2001)
Bosques tropicales	-2,19	2,54	0,98	105	>10	<3000mm	Chave <i>et al.</i> (2001) ^b
Bosques tropicales	-1,996	2,32	0,89	28	5-40	900-1500mm	Brown (1997)
Bosques tropicales	-2,134	2,530	0,97	170	5-148	1500-4000mm	Brown (1997)
Colombia	-1,97	2,248	0,97	54	8-98	MHT-	Overaman <i>et al.</i> (1990)
Cambodia	-2,39	2,56	0,99	22	--	--	Brown (1997) ^a
Caquetá, Colombia	-2,305	2,291	0,98	140	2-130	HT	Álvarez (1993)
Costa Rica	-1,81	2,32	0,98	92	--	--	Brown (1997) ^a
El Verde, Puerto Rico	-2,749	2,634	0,96	33	>4	HT	Crow (1980)
Guayana francesa	-2,88	2,72	0,97	1412	d	--	Chave <i>et al.</i> (2001)
Indonesia	-2,26	2,60	0,98	38	--	--	Brown (1997) ^a
Karnataka, India	-0,435	2,12	0,92	189	8-61	LT	Rai & Proctor (1986) ^d
Manaus, Brasil, Amazonía	-1,754	2,665	0,92	244	5-20	HTI	Higuchi <i>et al.</i> (1998)
Manaus, Brasil, Amazonía	-1,497	2,548	0,97	315	5-120	HT	Higuchi <i>et al.</i> (1998)
Manaus, Brasil, Amazonía	-0,151	2,170	0,90	28	>20	HT	Higuchi <i>et al.</i> (1998)
Nueva Guinea	-2,0	2,36	0,96	39	--	--	Edwards & Grubb (1977)
Para Brasil	-2,26	2,66	0,97	15	--	--	Brown (1997) ^a
Para Brasil	-2,55	2,65	0,97	127	--	--	Chave <i>et al.</i> (2001) ^e
Puerto Rico	-2,41	2,41	0,95	15	--	--	Chave <i>et al.</i> (2001)

^a Información publicada en Brown (1997).

^b Árboles ≥5cm.

^c Sólo 418 árboles ≥5cm.

^d Este modelo incluye un factor de corrección por sesgos propuesto por Mountford & Bunce (1973).

^e Estimado.

El valor de los parámetros estimados del modelo alométrico obtenido en este estudio (Tabla 2), se encuentran dentro del rango de valores estimados para los mismos parámetros en otros bosques tropicales como los de la Tabla 5. Para los diferentes bosques el parámetro b , de un modelo como (2), se mantiene muy constante y oscilando entre 2,12 y 2,66.

Por otra parte, al evaluar la precisión de este modelo con el índice \mathbf{dB} , se puede concluir que un sólo modelo no podría explicar la biomasa en todos los rangos diamétricos, sin cometer desviaciones superiores al 5% como lo aseveran Kira & Ogawa (1971). Álvarez (1993) atribuye estas imprecisiones de los modelos en las estimaciones, especialmente en las clases de tamaño superiores, a que probablemente la tasa alométrica de la relación biomasa y atributos de tamaño (como el diámetro), no permanece constante para todas las categorías, aún menos en rangos de tamaños tan amplios como los analizados aquí.

La aplicación de los factores de corrección (2) a (4) al modelo ajustado (Tabla 3) produce un desplazamiento de la línea ajustada, explicando que los valores de \mathbf{dB} aumenten y lo hagan a valores similares en cada factor. Beauchamp & Olson (1973) también encontraron gran similitud entre el factor (2) usado aquí y el propuesto por ellos; pero en ninguna referencia se ha encontrado comparaciones con el factor propuesto por Finney en 1941. El factor de corrección de Finney converge a los resultados de los factores (2) y (3), posiblemente debido a valores grandes de CME y al tamaño de la muestra (según Finney un valor grade de N es 50).

Estos resultados concuerdan con las consideraciones de Whittaker & Marks (1975) en el sentido de que al hacer las correcciones se pierde precisión en las estimaciones. Sin embargo, también se encuentran resultados opuestos. Por ejemplo, Jordan & Uhl (1978) encontraron 1,7% de subestimación de la biomasa aérea cuando utilizan el modelo corregido con la propuesta de Beauchamp & Olson (1973).

Independientemente de la precisión del modelo los factores de corrección conducen a resultados similares de predicción. Esta conclusión le da cierta ventaja al factor de corrección (2) por su sencillez y por la posibilidad de obtener una nueva representación del modelo insesgado, pero tiene la desventaja de que las pruebas convencionales realizadas para los parámetros no son válidas. Para los modelos de regresión con transformaciones logarítmicas se pueden desarrollar pruebas, como por ejemplo que el \hat{a}_1 es el mejor estimador insesgado de β_1 , aún sin cumplirse el supuesto de normalidad.

Esto es, tiene validez la prueba de t para el parámetro β . Pero para el parámetro β_0 es imposible realizar la prueba de t (Heien 1968); más aún después de sumarle el factor de corrección.

Una desventaja del factor de corrección propuesto por Beauchamp & Olson (1973) es que no está definido para modelos de regresión múltiple.

Existen trabajos que validan la omisión de los factores de corrección. Madgwick & Satoo (1975) hacen una presentación de varios resultados, comparando las estimaciones de biomasa de rodales tanto por el método de cosecha como con modelos de regresión. Concluyen que los modelos de regresión sin corrección por sesgos producidos por las transformaciones, tienden a sobreestimar la biomasa de las hojas entre 1% y 9% en los rodales. Estos autores simulan las estimaciones de biomasa en rodales de plantaciones, usando modelos de regresión generados con datos de muestreos estratificados por el diámetro de los árboles y con toda la muestra. El resultado es que con los muestreos estratificados se pueden obtener estimaciones del rodal desde 55% de subestimación hasta 80% de sobrestimación, con intervalos de confianza menos precisos que los obtenidos con toda la muestra. Se concluye que para estos rodales las estimaciones de la masa del fuste están dentro del 20% de error el 98% de las veces, mientras las ramas y las hojas están dentro del 20% del error el 53% y el 99% de las veces, respectivamente. Con estas variaciones los pequeños sesgos en las estimaciones de biomasa con los modelos estimados con transformaciones logarítmicas son de menor importancia.

Whittaker & Marks (1975) consideran que las correcciones por sesgos probablemente no mejoren mucho las estimaciones de la biomasa del bosque y de la productividad. Satoo & Madgwick (1982) consideran que las sobrestimaciones individuales producidas por los modelos son más serias que las subestimaciones. Por otra parte, Sprugel (1983) argumenta que si bien es cierto que los errores involucrados en las estimaciones por estos modelos son bajos (generalmente 10% o menos), se pueden calcular fácilmente y se deben corregir. Aunque otros factores pueden producir grandes e incorregibles errores dentro del análisis de regresión, no es excusa para desistir de la corrección, pues es una de las pocas fuentes de error evitables.

Orrego & del Valle (2001) han encontrado que en el área de Porce el compartimiento de biomasa aérea, representado por árboles y arbustos dicotiledóneos, participa con el 89% de la biomasa viva aérea total (237,255 t/ha) confirmando la importancia de este compartimiento en monitoreos de carbono. Si se asume que la información del inventario del año 2001 en 32 parcelas permanentes del área de Porce es parte de un proyecto de créditos de carbono, la decisión de omitir la corrección por sesgos al modelo de biomasa aérea implica subestimar las existencias de carbono en 15.373,90 (según el factor (2) de la Tabla 4) en un área de 609 ha. Dada la similitud en las correcciones por los

diferentes factores, se recomienda utilizar el factor (2) para corregir los sesgos de estimaciones de modelos de biomasa alométricos, ajustados por regresión lineal luego de hacer las transformaciones logarítmicas a las respectivas variables.

Agradecimientos

Esta investigación se realizó dentro de la línea de investigación Bosques y Cambio Climático, proyecto “Eficiencia de las Coberturas Vegetales en la Asimilación de CO₂, Central Hidroeléctrica Porce II”, dirigido por los profesores Jorge Ignacio del Valle y Sergio Alonso Orrego, cofinanciado por las Empresas Públicas de Medellín E.S.P. y la Universidad Nacional de Colombia, Sede Medellín, Contrato 3/DJ-1367/17, Acta 19. Queremos agradecer a H. Orozco, G. J. López de EPM y G. Vásquez de la Universidad Nacional por su apoyo logístico. Igualmente, a J. Manjarrés, M. Ocampo, H. Sejín, C. Sierra y M. A. Herrera.

Literatura citada

- ÁLVAREZ, D. E. 1993. Composición florística, diversidad, estructura y biomasa de un bosque inundable en la Amazonia Colombiana. Tesis Universidad de Antioquia Maestría en Biología. Facultad de Ciencias Exactas y Naturales, Medellín.
- BASKERVILLE, G. L. 1965. Estimation of dry weight of tree components and total stand crop in conifer stands. *Ecology* 46: 867-869.
- BEAUCHAMP, J. J., AND J. S. OLSON. 1973. Corrections for bias in regression estimates after logarithmic transformation. *Ecology* 54: 1403-1407.
- BELL, F. J., A. R. EK, H. C. HITCHCOCK, K. LES, C. I. MILLER, AND J. W. MILLER. 1984. Timber measurements. In F. K. Wenger (Ed.). *Forestry Handbook*, pp.314-315. John Wiley & Sons, New York.
- BRADU, D., AND Y. MUNDLAK. 1970. Estimation in lognormal linear models. *Journal of the American Statistical Association, Theory and Methods* 65: 198-211.
- BROWN, S. 1997. Estimating biomass and biomass change of tropical forest: a primer. *FAO Forestry Paper* 134, Roma.
- BROWN, S., AND A. E. LUGO. 1992. Aboveground biomass estimates for tropical moist forests of the Brazilian Amazon. *Interciencia* 17: 8-18.
- BRUNIG, E. F. 1983. Vegetation structure and growth. In F. B. Golley (Ed.). *Ecosystems of the world: tropical rain forest ecosystems structure and function*, pp.67-70. Elsevier Scientific Publishing Company, Amsterdam.

- CARROLL, J. R., AND D. RUPPERT. 1988. Transformation and weighting in regression. Chapman and Hall, New York.
- CHAVE, J., B. RIÉRA, AND M. DUBOIS. 2001. Estimation of biomass in a neotropical forest of French Guiana: spatial and temporal variability. *Journal of Tropical Ecology* 17:79-96.
- CROW, T. R. 1980. A rainforest chronicle: a 30-year record of change in structure and composition at El Verde, Puerto Rico. *Biotropica* 12: 42-55.
- DELANEY, M., S. BROWN, A. E. LUGO, A. TORRES-LEZAMA, AND N. B. QUINTERO. 1997. The distribution of organic carbon in major components of forests located in five life zones of Venezuela. *Journal of Tropical Ecology* 13: 697-708.
- EDWARDS, J. P, AND J. P. GRUBB. 1977. Studies of mineral cycling in a montane rain forest in New Guinea, I. The distribution of organic matter in the vegetation and soil. *Journal of Ecology* 65: 943-969.
- EMPRESAS PÚBLICAS DE MEDELLÍN. 1992. Proyecto Hidroeléctrico Porce II. Estudio de impacto ambiental. Informe Final. INTEGRAL, Medellín.
- ENQUIST, J. B., J. H. BROWN, AND G. B. WEST. 1998. Allometric scaling of plant energetic and population density. *Nature* 395:163-166.
- GREEN, J. E., AND H. T. VALENTINE. 1998. Bayesian analysis of the linear model with heterogeneous variance. *Forest Science* 44: 134-138.
- HEIEN, D. M. 1968. A note on log-linear regression. *Journal American Statistical Association* 63: 1034-1038.
- HIGUCHI, N., J. SANTOS, R. J. RIBEIRO, L. MIENTE, AND Y. BIOT. 1998. Biomassa da parte aérea da vegetação da floresta tropical úmida de Terra-Firme da amazônia brasileira. *Acta Amazonica* 28: 153-166.
- JORDAN, F. C., AND C. UHL. 1978. Biomass of a "tierra firme" forest of the Amazon basin. *Ecologia Plantarum* 13: 387-400.
- JUDGE, C. G., R. C. HILL, W. E. GRIFFITHS, H. LUTKEPOHL, AND T. LEE. 1987. Introduction to the theory and practice of econometrics. John Wiley & Sons. New York.
- KIRA, T., AND H. OGAWA. 1971. Assesment of primary production in tropical and equatorial forests, pp 309-321. *In* P. Duvigneaud (Ed.). Productivity of forest ecosystems. proceeding of the Brussels Symposium. UNESCO, Geneva.
- KITTREDGE, J. 1944. Estimation of the amount of foliage of trees and stand. *Journal of Forestry* 42: 905-912.
- LAURENT, A. G. 1963. The lognormal distribution and the translation method: description and estimation problems. *Journal of American Statistical Association* 58: 231-235.
- MACDICKEN, G. K. 1997. A guide to monitoring carbon storage in forestry and agroforestry projects. Winrock International Institute for Agricultural Development, USA.
- MADGWICK, H. A. I., AND T. SATOO . 1975. On estimating the aboveground weights of tree stands. *Ecology* 56: 1446-1450.

- MORRISSEY, A. W., AND J. R. JUSTUS. 1998. Global climate change. Committee for the National Institute for the Environment, Washington (Cnie@cnie.org).
- MOUNTFORD, M. D., AND R. G. H. BUNCE 1973. Regression sampling with allometrically related variables, with particular reference to production studies. *Forestry* 46: 203-212.
- ORREGO, S. Y J. I. DEL VALLE. 2001. Eficiencia de las coberturas vegetales en la asimilación de CO₂. Informe Final. Empresas Públicas de Medellín E.S.P., Universidad Nacional de Colombia, Sede Medellín, Medellín.
- OVERMAN, M. P. J., J. G. SALDARRIAGA, Y J. F. DUIVENVOORDEN. 1990. Estimación de la biomasa aérea en el bosque del medio Caquetá, Colombia. *Colombia Amazonica* 4: 135-147.
- _____, HENDRIK J. W., AND J. G. SALDARRIAGA. 1994. Evaluation of regression models for aboveground biomass determination in Amazon rainforest. *Journal of Tropical Ecology* 10: 207-218.
- RAI, S. N., AND J. PROCTOR. 1986. Ecological studies on four rainforests in Karnataka, India I. Environment, structure, floristic and biomass. *Journal of Ecology* 74: 439-454.
- REGINA, S. I. 2000. Biomass estimation and nutrient pools in four *Quercus pyrenaica* in Sierra de Gata, Salamanca, Spain. *Forest Ecology and Management* 132: 127-141.
- SALDARRIAGA, J. G., D. C. WEST, M. L. THARP, AND C. UHL, 1988. Long-term chronosequence of forest succession in the upper Río Negro of Colombia and Venezuela. *Journal of Ecology* 76: 938-958.
- SATOO, T., AND H. A. I. MADGWICK 1982. Forest biomass. Martinus Nijhoff/Drw. Junk Publishers, Netherlands.
- SCATENA, F. N., W. SILVER, T. SICCAMA, A. JONSON, AND M. J. SÁNCHEZ. 1993. Biomass and nutrient content of the Bisley experimental watersheds, Luquillo Experimental Forest, Puerto Rico, before and after hurricane Hugo, 1989. *Biotropica* 25: 15-27.
- SPRUGEL, D. G. 1983. Correcting for bias in log-transformed allometric equations. *Ecology* 64: 209-210.
- THÖNI, H. 1969. A table for estimating the mean of a lognormal distribution. *Journal of the American Statistical Association* 64: 632-636.
- WHITTAKER, R. H., AND L. P. MARKS. 1975. Methods of assessing terrestrial productivity, pp 113-124. In H. Leith and R. H. Whittaker (Eds.). Primary productivity of the biosphere. Ecological Studies. 14, Springer Verlag, New York.
- ZAPATA, M. 2001. Ecuaciones de biomasa aérea para los bosques primarios del área de influencia de la Central Hidroeléctrica Porce II. Trabajo de grado para optar al título de Ingeniero Forestal, Universidad Nacional de Colombia, Sede Medellín, Medellín.